DOI:
10.1136/jmedgenet-2018-105726

Document Version
Peer reviewed version

Genome-wide meta-analysis identifies genetic locus on chromosome 9 associated with Modic changes

Maxim B Freidin, PhD[1]*, Minna Kraatari, MD[2,3,4]*, Sini Skarp, MSc[2,3]*, Juhani Määttä, MD, PhD[1,4], Johannes Kettunen, PhD[2,5,6], Jaakko Niinimäki[4,7], MD, PhD, Jaro Karppinen, MD, PhD[2,4,8], Minna Männikkö, PhD[9+], Frances MK Williams, PhD, FRCP(E)[1+]

[1]*Department of Twin Research and Genetic Epidemiology, School of Life Course Sciences, King's College London, London, UK.*

[2]*Center for Life Course Health Research, Faculty of Medicine, University of Oulu, Finland*

[3]*Faculty of Biochemistry and Molecular Medicine, University of Oulu, Finland*

[4]*Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Finland*

[5]*Computational Medicine, Faculty of Medicine, University of Oulu and Biocenter Oulu, Oulu, Finland*

[6] *National Institute for Health and Welfare, Helsinki, Finland*

[7] *Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Oulu, Finland*

[8]*Finnish Institute of Occupational Health, Oulu, Finland*

[9]*Northern Finland Birth Cohorts, Faculty of Medicine, University of Oulu, Oulu, Finland*

*These authors contributed equally

[+]Shared last authorship

Corresponding author: Minna Männikkö, Northern Finland Birth Cohort Studies, Faculty of Medicine, Aapistie 5 / P.O. Box 5000, 90014 University of Oulu, Finland, minna.ruddock@oulu.fi, +358-(0) 294 -485751.

Word count: 3223

ABSTRACT

Background: Low back pain (LBP) is a common disabling condition. Lumbar disc degeneration (LDD) may be a contributing factor for LBP. Modic change (MC), a distinct phenotype of LDD, is presented as a pathological bone marrow signal change adjacent to vertebral endplate on magnetic resonance imaging (MRI). It is strongly associated with LBP and has heritability around 30%. Our objective was to identify genetic loci associated with MC using a genome-wide meta-analysis.

Methods: Presence of MC was evaluated in lumbar MRI in the Northern Finland Birth Cohort 1966 (N=1182) and TwinsUK (N=647). Genome-wide association analyses were carried out using linear regression model. Inverse-variance weighting approach was used in themeta-analysis.

Results: A locus associated with MC (p<5e-8) was found on chromosome 9 with the lead SNP rs1934268 in an intron of the *PTPRD* gene. It is located in the binding region of BCL11A, SPI1 and PBX3 transcription factors. The SNP was nominally associated with LBP in TwinsUK (p = 0.001), but not associated in the UK Biobank (p = 0.914). Suggestive signals (p<1e-5) were identified near *XKR4*, *SCIN*, *MGMT*, *DLG2*, *ZNF184*, and *OPRK1*.

Conclusion: *PTPRD* is a novel candidate gene for MC that may act via the development of cartilage or nervous system; further work is needed to define the mechanisms underlying the pathways leading to development of MC. This is the first genome-wide meta-analysis of MC and the results pave the way for further studies on the genetic factors underlying the various features of spine degeneration and LBP.

INTRODUCTION

Low back pain (LBP) is highly prevalent and disabling musculoskeletal condition throughout the world causing significant burden to the society [1]. In Europe, it is estimated to consume as much of a country's GDP as 0.7% - 2.2%. Lumbar disc degeneration (LDD) is likely a risk factor for LBP,

based on multiple epidemiologic studies [2-6], even though it is not universally accepted due to the presence of confirmed disc pathology among asymptomatic individuals [7] and the presence of LBP among patients without LDD [8]. The role of genetic factors in the development of LDD has been demonstrated to be significant, with heritability estimates ranging from 64 to 81% [9].

Modic change (MC), subchondral and vertebral bone marrow change visible on spine magnetic resonance (MR) imaging, are reported in 20%-50% of patients with LDD [10-12]. MC can be divided into three different types according to T1- and T2-weighted MRIs: type 1 (MC1) (low T1 and high T2 signals) indicating an ongoing degeneration, type 2 (MC2) (high T1 and high T2 signals) indicating fatty replacement of the bone marrow and type 3 (MC3) (low T1 and low T2 signals) indicating bone sclerosis [10-12]. MC can convert from one type to another and also mixed types have been identified [11-13]. MC has been strongly associated with LBP (odds ratio, OR = 6.1 (95% confidence interval, CI = 2.9–13.1)) and poor reduction of pain (β = −9.79 (95% CI = −16.7 to −2.90)) in clinical groups [14,15] and with LBP in general population samples (OR ranging from 1.9 (95% CI = 1.2–3.0) to 4.5 (95% CI = 1.7–11.6)) [16]. Like LDD, MC is common with the prevalence of MC up to 56% in general population [17,18] and 81% in clinical samples [19,20].

The genetic background of MC is largely unknown but the heritability of MC has been estimated in the range 16%-43% [21]. Previously, the genetic basis of MC has been studied using candidate gene approach and whole exome sequencing. Interleukin-1α (*IL1A*) and matrix metallopeptidase 3 (*MMP3*) were found to be significantly associated with MC (OR = 3.2 (95% CI = 1.2-8.5); OR=2.50 (95% CI = 1.09–5.71, respectively) [22,23] and variants in mastermind like transcriptional coactivator 1 (*MAML1*) and heparan sulfate proteoglycan 2 (*HSPG2*) genes co-segregated with MC[24]. Differential expression of tumor necrosis factor α (TNF-α), ADAM metallopeptidase with thrombospondin type 1 motif 5 (*ADAMTS5*), and NLRP3/caspase-1/interleukin-1β axis has been reported in cartilaginous endplate of patients with MC [25,26].

Given that the pathogenic mechanisms underlying MC are unknown but genetic factors do play a role, agnostic approaches such as genome-wide association study (GWAS), may provide information about underlying genetically-mediated pathways of pathology. Herewith we report the first ever GWAS for MC based on the meta-analysis of two population samples of Northern European adults.

METHODS

Study samples

Two population samples have been used in the current study: TwinsUK and Northern Finland Birth Cohort 1966 (NFBC1966).

The TwinsUK register comprises female twin pairs recruited from the general UK population since 1992 [27]. The twins have attended for lumbar MRI, nurse-led interview and collection of demographic data between 1996 and 2000 and lumbar MRI was performed for a subsample a decade later. The interview used a modified version of the UK Medical Research Council Back and Neck pain questionnaire and assessed a lifetime history of low back and neck pain, as well as psychological distress, lifestyle variables (e.g. physical exercise) and socioeconomic status, as described elsewhere [6]. Twins gave informed written consent and appropriate ethics permission had been obtained from the St Thomas' Hospital Research Ethics Committee. Genotyping and imputation of the TwinsUK dataset was carried out as described previously [28]. Imputation was performed using 1000 Genomes phase 3 reference panel. The sample available for use in the current study had MC coded on MRI comprised 647 individuals including 625 females and 22 males. The sample mean age was 54.7±8.9 years and mean BMI was 25.2±4.7 kgm^{-2}.

NFBC1966 comprises children born in Oulu and Lapland (Finland) who had expected dates of birth between Jan 1st and Dec 31st, 1966 (12,068 mothers, 12,231 children). Participants' health, medication, healthcare utilisation, lifestyle habits and social status have been collected at ages 14, 31

and 46 years by questionnaire and a wide-scale health examination performed at 46 years. A subset of study subjects that had attended the examinations underwent lumbar MRI (n=1540). A total of 5400 individuals were chosen for genotyping which was performed using Illumina Human CNV370-Duo DNA bead chip as described previously [29]. The Ethical Committee of the Northern Ostrobothnia Hospital District has approved the study and the study subjects have given their informed written consent. The sample used for the current study comprised 1182 individuals including 642 females and 540 males. The mean age of the sample was 47.4±0.7 years and the mean BMI was 26.6±4.6 kgm^{-2}.

Magnetic resonance imaging and coding for MC

In TwinsUK, the lumbar MRI at baseline and follow-up was performed using a 1.0-T equipment (Siemens, Munich, Germany) as previously reported [9]. In T2-weighted images a fast spin-echo sequence of time to recovery (TR) 5000–4500 ms and time to echo (TE) 112 ms, with a slice thickness of 4 mm were used.

In NFBC1966, MRI was performed using a 1.5-T imaging system (Signa HDxt, General Electric, Milwaukee, Wisconsin, USA). The routine lumbar spine protocol was used with T2-weighted fast-recovery fast spin-echo images in transverse and sagittal planes. In transverse planes time TE 118 ms and TR 4000 ms with slice thickness of 4 mm was used and in sagittal planes TE 112 ms and TR 5000 ms with slice thickness of 3 mm was used.

Coding for MC was carried out using standard methods in both TwinsUK and NFBC1966 by, or under the oversight of, co-author JM. In TwinsUK, a randomly chosen subset of the scans was re-coded by the second observer (n=50). In NFBC1966, two observers coded all available scans in half and a subset of scans (n=66) was coded by both observers for the calculation of inter-rater reliability. Both rostral and caudal endplates were examined and the presence, type, height and width of MC

recorded as described previously [21]. MC was coded as type 1, type 1/2, type 2, type 2/3 and type 3 in NFBC1966. In TwinsUK, the type of MC could not be fully assessed as only T2-weighted images had been obtained. In both cohorts, MC was measured by evaluating the maximum height and width of MC in rostral and caudal endplates. The maximum height of each MC was compared with the height of the vertebral body and coded as the following: 0) no MC; 1) MC in the endplate only; 2) MC height is less than 25% of the vertebral body; 3) MC height ranging from 25% to 50% of the vertebral body; and 4) MC height is more than 50% of the height of the vertebral body. The horizontal width of each MC was graded by dividing the vertebral body axially into four peripheral quadrants and one central circular zone with width ranging from 0 to 5. The mean (±SD) of the height grade of MC was 2.2±SD in NFBC1966 and 1.9±0.7 in TwinsUK, while the mean of the width grade of MC was 2.8±SD in NFBC1966 and 2.5±1.2 in TwinsUK. The inter-rater reliability for the presence of MC was kappa = 0.92 in TwinsUK; and kappa = 0.82 in NFBC1966.

Genome-wide association study and meta-analysis

For the GWAS, we did not stratify or filter out the subjects by the grade of MC; thus, the outcome phenotype was a categorical trait defined as the presence of MC regardless of the type or size of MC vs the absence of MC. Accordingly, the phenotype was coded as 1 ("MC of any type or size") or 0 ("no MC"). For NFBC1966, association analysis was carried out using SNPTEST v2.5.2, while for TwinsUK it was carried out using GEMMA v0.9. A linear regression model was fitted to test for additive effects of SNPs (genotype dosage) adjusting for age, sex, BMI, and either family relatedness via a kinship matrix (for TwinsUK) or population stratification using genotypic principal components (for NFBC1966). The use of linear regression for a categorical trait was dictated by the need to apply linear mixed-effects model for TwinsUK to account for family structure and the lack of statistical software that would allow such the analysis using logistic regression. Given that linear regression effect sizes are not straightforward to interpret in terms of the risks for categorical traits, we calculated

odds ratios (OR) and corresponding 95% confidence intervals (CI) using the following approximation: OR = exponent(β / (μ × (1 - μ))), where β is the linear effect size and μ is the case fraction in the dataset.

Meta-analysis of the results obtained from NFBC1966 and TwinsUK samples was carried out by inverse-variance weighting approach using METAL software[30]. Prior to meta-analysis SNPs in both datasets were filtered out using the following settings: monomorphic SNP excluded; minor allele frequency < 0.03; deviation from Hardy-Weinberg equilibrium p-value < 1e-6 (to minimize the amount of SNPs with likely erroneous genotyping); genotyping call rate < 0.95; and imputation quality < 0.7. For NFBC1966 dataset, additionally we excluded SNPs for which model information estimated by SNPTEST was < 0.7. The total number of SNPs that were present in both datasets after applying the filters was 5,479,070.

Genome-wide significance threshold was p < 5e-8; suggestive association thresholds were p < 5e-7 and p < 1e-5. Post-GWAS analysis involved look-ups in the public databases including Phenoscanner [31], GTEx [32], LDlink [33], and RegulomeDB [34].

In addition, we determined if SNPs associated with MC are also associated with chronic LBP (considered as pain between lowest rib and gluteal fold lasting 3 months or longer) in TwinsUK dataset (n = 4267) and acute or chronic BP ("Back pain" response to a question: "Pain type(s) experienced in last months") using our ongoing study of UK Biobank (n = 453,862).

RESULTS

A total of 1829 individuals from two cohorts of Northern European ancestry (Finland and the UK) were examined by a GWAS meta-analysis. Overall, there were 994 individuals having MC detected by MRI (257 in TwinsUK and 737 in NFBC1966 cohorts) and 835 without MC (390 in TwinsUK and 445 in NFBC1966 cohorts).

In TwinsUK, a locus on chromosome 4 upstream *SPOCK3* gene was found to be associated with MC, with the lead SNP rs72703315 (OR = 0.25, 95% CI = 0.15-0.40; p = 2.04e-08) (Supplementary Figure 1, Supplementary Figure 2). In NFBC1966, a locus on chromosome 9 in intron of *PTPRD* gene was found to be associated with MC, with the lead SNP rs1934268 (OR = 1.79, 95% CI = 1.48-2.18; p = 7.28e-09) (Supplementary Figure 3, Supplementary Figure 4). In both cohorts the test-statistic inflation (lambda) for the GWAS was 1.03 suggesting no significant population stratification.

The results for SNPs with genome-wide significance (p<5e-8) and suggestive significance for association (p<5e-7) are listed in Table 1. Meta-analysis revealed a single locus on chromosome 9 associated with MC with a genome-wide significance (Table 1, Figure 1, Figure 2). The test statistic inflation lambda = 1.02, suggesting no population stratification. The lead SNP rs1934268 was the same as that identified in the NFBC1966 cohort, located in an intron of the *PTPRD* gene which encodes a protein tyrosine phosphatase receptor type D (Figure 3). There was moderate heterogeneity between the studies for the lead SNP and suggestive SNPs in the region, though not statistically significant (p>0.05). The effect size of the association was rather high (for the lead SNP rs1934268, OR = 1.59, 95% CI = 1.36-1.86) suggesting a 59% increase in risk of MC for rs1934268*T allele vs rs1934268*C allele. This value was approximated from linear regression effect size so may be an overestimate. According to Phenoscanner, no associations have been reported for any other trait so far for the lead SNP rs1934268 or those in strong LD with it (r2>0.7).

Table 1. The results of meta-GWAS for Modic changes in TwinsUK and NFBC1966 cohorts.

SNP ID	Chr:Pos (hg19)	Effect allele	Other allele	EAF ± SE	Beta ± SE	OR (95% CI)	P-value	Direction	I^2	Het. P-value
rs1934268	9:9730366	T	C	0.788 ± 0.032	0.115 ± 0.020	1.59 (1.36-1.86)	7. 4E-09	++	66.5	0.084
rs4382526	9:9726395	A	G	0.206 ± 0.031	-0.105 ± 0.020	0.66 (0.56-0.77)	1.2E-07	--	62.2	0.104
rs7047812	9:9726090	A	C	0.206 ± 0.031	-0.105 ± 0.020	0.66 (0.56-0.77)	1.3E-07	--	62.5	0.102
rs7862707	9:9723463	A	G	0.206 ± 0.030	-0.105 ± 0.020	0.66 (0.56-0.77)	1.4E-07	--	62.9	0.100
rs5896346	9:9708238	T	TC	0.793 ± 0.030	0.105 ± 0.020	1.53 (1.30-1.79)	1.6E-07	++	57.2	0.126

Chr:pos, chromosome:position; EAF, effect allele frequency; Beta, coefficient of linear regression for effect allele vs other allele; OR, Odds ratio calculated using the following approximation: OR = $\exp(\beta/(\mu\times(1-\mu))$, where μ is the prevalence of MC (0.54 in the joint dataset); I^2, heterogeneity statistics.

According to RegulomeDB, the rs1934268 SNP is located in the binding region for transcription factors BCL11A, SPI1 and PBX3 (Supplementary Figure 5), although the potential of this variant to affect binding is considered minimal (Regulome DB score = 5). The SNP is not known as an eQTL according to GTEx database. Moreover, no SNPs in high LD (r^2>0.7) with the lead SNP are known to be functionally important according to LDlink and RegulomeDB. However, considering D' as a measure of LD, a rare SNP rs186163656 in perfect LD with rs1934268 (D' = 1) exhibits likely effect on the binding of CTCF transcription factor (RegulomeDB score = 2c) in an adjacent site.

None of the other SNPs, apart from those on chromosome 9 at the *PTPRD* locus, achieved a strict suggestive significance threshold of 5e-7. However, 34 SNPs had p-values < 1e-5 which is considered

a threshold for suggestive association (Supplementary Table). These SNPs overlap or locate closely to the genes *PTPRD*, *XKR4*, *SCIN*, *MGMT*, *DLG2*, *ZNF184*, and *OPRK1* with no heterogeneity between the studies in the majority of cases.

We investigated whether the lead SNP rs1932568 is associated with back pain. In TwinsUK, the SNP was nominally associated with chronic LBP (OR = 1.26, 95%CI = 1.10-1.44; p = 0.001); however, in the much larger sample from UK Biobank, this association was not observed (OR = 1.00, 95% CI = 0.51-1.96; p = 0.914).

DISCUSSION

Using meta-analysis of two cohorts (TwinsUK and NFBC1966) with the total sample size of 1829, we identified a locus on chromosome 9 significantly associated with MC. While the lead SNP rs1934268 in the intron of *PTPRD* gene has not previously been associated with any trait, the association region is likely functional with at least one SNP (rs186163656) in strong LD with the lead SNP being able to affect binding of transcription factors. Transcription factor binding sites at or adjacent to the rs1934268 locus include SPI1, PBX3, BCL11A and CTCF. SPI1 participates in regulation of osteoclast differentiation [35]. CTCF and PBX3 are involved in multiple pathways including regulation of homeodomain proteins during the development of embryonic anterior-posterior axis and limb buds [36,37], while BCL11A participates in spinal cord development [38]. These data suggest that the activity of *PTPRD* may be affected by genetic variants in the site via influence on the binding of a number of transcription factors during bone growth and embryonic development of the musculoskeletal system.

The *PTPRD* gene encodes protein tyrosine phosphatase receptor type D, a member of the protein tyrosine phosphatase (PTP) family that is involved in controlling a variety of cellular processes including cell growth, differentiation, mitotic cycle and oncogenic transformation. It is critically important for neural system development via promoting neurite growth and axon guidance [39] and

it also acts as a tumour suppressor [40]. According to GTEx database, the gene is highly expressed in brain tissues including spinal cord. More importantly in the context of MC, the gene is overexpressed in intact cartilage as compared with damaged osteoarthritic cartilage [41], and its expression is associated with clinical success after autologous chondrocyte implantation [42]. Furthermore, in mice, homologues *Ptprd* gene was found to be expressed in vertebral cartilage condensation during spinal development (http://www.informatics.jax.org/tissue/marker/MGI:97812). Also, the *PTPN11* gene, another member of the PTP family, was established to play a critical role in cartilage development and homeostasis [43]. Thus, *PTPRD* gene appears to be a functionally plausible candidate gene for the development of MC. Interestingly, in TwinsUK, the direction of the effect of the SNP on MC and chronic LBP are consistent with the T allele predisposing to the higher risk of both MC and chronic LBP. This is concordant with a hypothesis of MC as a risk factor for LBP and provides evidence of underlying molecular factors [44].

Of the suggestive associations identified in the current study (Supplementary table 1), *OPRK1* is of particular interest. Even though the signal for this gene was not seen independently in two participating cohorts, it is remarkably more significant in the meta-analysis and shows no heterogeneity between the studies, thus indicating a strong trend towards replication. According to the GTEx, all the SNPs near *OPRK1* with suggestive association with MC (Supplementary Table) are eQTLs for this gene. The protein encoded by the gene is an opioid receptor involved in pain perception and variation in its expression is observed in the lumbar spinal cord in neuropathic pain model [45]. Moreover, it is highly expressed in prechondrocytes and involved in protection against degradation in injured cartilaginous tissues [46].

To our knowledge, this is the first ever GWAS reported for MC. Identifying the genetic factors contributing to MC increases our knowledge of the etiology and pathology of this trait, which might in future lead to identification of novel biomarkers or therapeutic targets for LBP and intervertebral

disc degeneration. As a GWAS study, the results allow limited interpretation at this stage, owing to the lack of available functional data on human disc such as eQTL or chromatin features. However, if we are to move the field forward, a whole range of techniques, including genetic ones, will need to be brought to bear. These results are an important first step and that the genetic loci identified may offer a highly relevant pointer to mechanisms of pathogenesis in MC leading to LBP.

The study has limitations. The sample size was not high according to the state-of-the art requirements for GWAS; however, this was dictated by the availability of cohorts having costly spine MR imaging with MC coded using the same methods, as well as genome-wide genotypes in ethnically similar samples. Replication in other datasets is needed to validate our findings further. Furthermore, functional experimental studies are required to assess the impact of *PTPRD* on the development of MC. Both replication and functional validation are necessary to ensure that our finding is not false-positive. These issues are not specific to our study and are common for complex traits GWAS. It is apparent that the significant association in the meta-analysis was driven by NFBC1966 cohort and a single significant locus present in TwinsUK on chromosome 4 was not reproduced in the meta-analysis. There was also a heterogeneity between the cohorts. The prevalence of MC in TwinsUK and NFBC1966 was different (39.7% and 62.3%, respectively) likely reflecting samples differing in mean age and sex ratio, both factors being known to be contribute to the risk of MC [20]. Also, different MRI equipment and field-strengths were used in NFBC1966 and TwinsUK for unavoidable reasons. A 1.5-T equipment was used in NFBC1966 and 1.0-T equipment in TwinsUK. 1.5-T indicates that the field strength is 1.5 tesla whereas 1.0-T equipment has a magnetic field of 1.0 tesla. This theoretically could have impacted the MC evaluation - MC is more often observed in higher-field MRI [47]. This is unlikely, though, given that the difference in the magnetic field strengths is not great at 0.5T.

TwinsUK, despite being older on average than NFBC1966, exhibited lower prevalence of MC, so the higher number of males in NFBC1966 may explain the difference in the prevalence of MC. To

address this, we tested chromosome 4 and 9 loci for gender-specific effects in NFBC1966 (impractical in TwinsUK as males n = 22). In this stratified analysis, there was no association between MC and the lead SNP on chromosome 4 (OR = 0.99, 95% CI = 0.58-1.69, p = 0.9774; OR = 1.01, 95% CI = 0.62-1.65, p = 0.9647, in men and women respectively). However, there was a significant unidirectional association between MC and the lead SNP on chromosome 9 in males and females (OR = 1.96, 95% CI = 1.48-2.61, p = 2.13e-6; OR = 1.71, 95% CI = 1.30-2.25, p = 1.24e-4, respectively).

It seems unlikely, therefore, that the association with the chromosome 9 region is gender-specific, while the difference between the results in TwinsUK and NFBC1966 for chromosome 4 region may reflect other differences not measured between the cohorts, such as ethnic specificity in MC prevalence and gene-environmental interactions. While MC has been shown in both small clinical and large epidemiological studies to be an independent risk factor for LBP, we did not find any association between rs1934268 and BP in the very large UK Biobank sample. This suggests perhaps that in the general population represented by the UK Biobank participants, MC-mediated BP does not contribute a large enough genetic signal to be detectable amongst all the other causes of BP. Further work should, therefore, concentrate on the aetiology of MC and more specific subgroup GWAS analysis performed in chronic LBP.

In summary, our GWAS meta-analysis for MC provides evidence of genetic contribution to the development of this trait and identifies novel candidate genes. Further functional studies may provide insight into the genesis of this trait and, eventually, clarify the role of MC in LBP.

ACKNOWLEDGEMENTS

We wish to thank all the participants in the TwinsUK, NFBC 46y study and UK Biobank and the NFBC project center.

CONTRIBUTORS

MF, MK, SS, JKe, JKa, MM, FW designed the study. MK, JM, JN, JKa collected the data. MF, MK, SS performed the data analysis and drafted the manuscript. MM and FW coordinated the study. All authors commented on the manuscript and agreed on the final version to be published.

FUNDING

This study was supported by EU FP7 project Pain_OMICS (grant agreement #602736).

TwinsUK is funded by the Wellcome Trust, Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.

NFBC1966 data collection at 46y received financial support from the University of Oulu (Grant no. 24000692), Oulu University Hospital (Grant no. 24301140), and the European Regional Development Fund (Grant no. 539/2010 A31592).

The UK Biobank Resource data were used under the project "Genetic and epidemiological analyses of low back pain" (#18219).

COMPETING INTERESTS

None to declare

PATIENT CONSENT

Obtained

ETHICS APPROVAL

St Thomas' Hospital Research Ethics Committee and The Ethical Committee of the Northern Ostrobothnia Hospital District have approved the study.

REFERENCES

[1] Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basanez MG, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabe E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fevre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FG, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF, Gmel G, Gonzalez-Medina D, Gosselin R, Grainger R, Grant B, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo JP, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Laden F, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Levinson D, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora ME, Meltzer M, Mensah GA, Merriman TR, Meyer AC, Miglioli V, Miller M, Miller TR, Mitchell PB, Mock C, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KM, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk GV, Polinder S, Pope CA,3rd, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson C, De Leon FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJ, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiebe N, Wiersma ST, Wilkinson JD, Williams HC, Williams SR, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh PH, Zaidi AK, Zheng ZJ, Zonies D, Lopez AD. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2013;**380**:2197-2223.

[2] Livshits G, Popham M, Malkin I, Sambrook PN, Macgregor AJ, Spector T, Williams FM. Lumbar disc degeneration and genetic factors are the main risk factors for low back pain in women: the UK Twin Spine Study. *Ann Rheum Dis* 201;**70**:1740-1745.

[3] Brinjikji W, Diehn FE, Jarvik JG, Carr CM, Kallmes DF, Murad MH, Luetmer PH. MRI Findings of Disc Degeneration are More Prevalent in Adults with Low Back Pain than in Asymptomatic Controls: A Systematic Review and Meta-Analysis. *AJNR Am J Neuroradiol*;**36**:2394-2399.

[4] Jensen RK, Kent P, Jensen TS, Kjaer P. The association between subgroups of MRI findings identified with latent class analysis and low back pain in 40-year-old Danes. *BMC Musculoskelet Disord* 2018;**19**:62.

[5] Kjaer P, Leboeuf-Yde C, Sorensen JS, Bendix T. An epidemiologic study of MRI and low back pain in 13-year-old children. *Spine (Phila Pa 1976)* 2005;**30**:798-806.

[6] MacGregor AJ, Andrew T, Sambrook PN, Spector TD. Structural, psychological, and genetic influences on low back and neck pain: a study of adult female twins. *Arthritis Rheum* 2004;**51**:160-167.

[7] Brinjikji W, Luetmer PH, Comstock B, Bresnahan BW, Chen LE, Deyo RA, Halabi S, Turner JA, Avins AL, James K, Wald JT, Kallmes DF & Jarvik JG (2015b) Systematic literature review of imaging features of spinal degeneration in asymptomatic populations. AJNR Am J Neuroradiol 36(4): 811-816.

[8] Lateef H & Patel D (2009) What is the role of imaging in acute low back pain? Curr Rev Musculoskelet Med 2(2): 69-73.

[9] Sambrook PN, MacGregor AJ, Spector TD. Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. *Arthritis & Rheumatism* 1999;**42**:366-372.

[10] de Roos A, Kressel H, Spritzer C, Dalinka M. MR imaging of marrow changes adjacent to endplates in degenerative lumbar disk disease. *AJR Am J Roentgenol* 1987;**149**:531-534.

[11] Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. *Radiology* 1988;**166**:193-199.

[12] Modic MT, Masaryk TJ, Ross JS, Carter JR. Imaging of degenerative disk disease. *Radiology* 1988;**168**:177-186.

[13] Kuisma M, Karppinen J, Niinimäki J, Kurunlahti M, Haapea M, Vanharanta H, Tervonen O. A three-year follow-up of lumbar spine endplate (Modic) changes. *Spine* 2006;**31**:1714-1718.

[14] Albert HB, Manniche C. Modic changes following lumbar disc herniation. *Eur Spine J* 2007;**16**:977-982.

[15] Jensen OK, Nielsen CV, Sorensen JS, Stengaard-Pedersen K. Type 1 Modic changes was a significant risk factor for 1-year outcome in sick-listed low back pain patients: a nested cohort study using magnetic resonance imaging of the lumbar spine. *Spine J* 2014;**14**:2568-2581.

[16] Kjaer P, Leboeuf-Yde C, Korsholm L, Sorensen JS, Bendix T. Magnetic resonance imaging and low back pain in adults: a diagnostic imaging study of 40-year-old men and women. *Spine* 2005;**30**:1173-1180.

[17] Wang Y, Videman T, Battie MC. Modic changes: prevalence, distribution patterns, and association with age in white men. *Spine J* 2012;**12**:411-416.

[18] Mok FP, Samartzis D, Karppinen J, Fong DY, Luk KD, Cheung KM. Modic changes of the lumbar spine: prevalence, risk factors, and association with disc degeneration and low back pain in a large-scale population-based cohort. *Spine J* 2016;**16**:32-41.

[19] Jensen TS, Bendix T, Sorensen JS, Manniche C, Korsholm L, Kjaer P. Characteristics and natural course of vertebral endplate signal (Modic) changes in the Danish general population. *BMC Musculoskelet Disord* 2009;**10**:81.

[20] Arana E, Kovacs FM, Royuela A, Estremera A, Asenjo B, Sarasíbar H, Amengual G, Galarraga I, Alonso A, Casillas C, Muriel A, Montoya J, Ordóñez C, Martínez C, Zamora J, Campillo C, Abraira V. Modic changes and associated features in Southern European chronic low back pain patients. *The Spine Journal* 2011;**11**:402-411.

[21] Määttä J, Kraatari M, Wolber L, Niinimäki J, Wadge S, Karppinen J, Williams F. Vertebral endplate change as a feature of intervertebral disc degeneration: a heritability study. *European Spine Journal* 2014;**23**:1856-1862.

[22] Karppinen J, Solovieva S, Luoma K, Raininko R, Leino-Arjas P, Riihimaki H. Modic changes and interleukin 1 gene locus polymorphisms in occupational cohort of middle-aged men. *Eur Spine J* 2009;**18**:1963-1970.

[23] Karppinen J, Daavittila I, Solovieva S, Kuisma M, Taimela S, Natri A, Haapea M, Korpelainen R, Niinimäki J, Tervonen O, Ala-Kokko L, Mannikko M. Genetic factors are associated with modic changes in endplates of lumbar vertebral bodies. *Spine (Phila Pa 1976)* 2008;**33**:1236-1241.

[24] Kraatari M, Skarp S, Niinimäki J, Karppinen J, Männikkö M. A Whole Exome Study Identifies Novel Candidate Genes for Vertebral Bone Marrow Signal Changes (Modic Changes). *Spine (Phila Pa 1976)* 2017;**42**:1201-1206.

[25] Chen S, Huang Y, Zhou ZJ, Hu ZJ, Wang JY, Xu WB, Fang XQ, Fan SW. Upregulation of tumor necrosis factor alpha and ADAMTS-5, but not ADAMTS-4, in human intervertebral cartilage endplate with modic changes. *Spine (Phila Pa 1976)* 2014;**39**:817-25.

[26] Tang P, Zhu R, Ji WP, Wang JY, Chen S, Fan SW, Hu ZJ. The NLRP3/Caspase-1/Interleukin-1beta Axis Is Active in Human Lumbar Cartilaginous Endplate Degeneration. *Clin Orthop Relat Res* 2016;**474**:1818-1826.

[27] Moayyeri A, Hammond CJ, Hart DJ, Spector TD. The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet* 2013;**16**:144-149.

[28] Ruth KS, Campbell PJ, Chew S, Lim EM, Hadlow N, Stuckey BG, Brown SJ, Feenstra B, Joseph J, Surdulescu GL, Zheng HF, Richards JB, Murray A, Spector TD, Wilson SG, Perry JR. Genome-

wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. *Eur J Hum Genet* 2016;**24**:284-290.

[29] Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, Jones CG, Zaitlen NA, Varilo T, Kaakinen M, Sovio U, Ruokonen A, Laitinen J, Jakkula E, Coin L, Hoggart C, Collins A, Turunen H, Gabriel S, Elliot P, McCarthy MI, Daly MJ, Jarvelin MR, Freimer NB & Peltonen L. *Nat Genet* 2009;**41**:35-46.

[30] Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;**26**:2190-2191.

[31] Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, Paul DS, Freitag D, Burgess S, Danesh J, Young R, Butterworth AS. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics* 2016;**32**:3207-3209.

[32] GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;**348**:648-660.

[33] Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015;**31**:3555-3557.

[34] Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;**22**:1790-1797.

[35] Tondravi MM, McKercher SR, Anderson K, Erdmann JM, Quiroz M, Maki R, Teitelbaum SL. Osteopetrosis in mice lacking haematopoietic transcription factor PU.1. *Nature* 1997;**386**:81-84.

[36] Capellini TD, Zappavigna V, Selleri L. Pbx homeodomain proteins: TALEnted regulators of limb patterning and outgrowth. *Dev Dyn* 2011;**240**:1063-1086.

[37] Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D. Functional analysis of CTCF during mammalian limb development. *Dev Cell* 2010;**19**:819-830.

[38] John A, Brylka H, Wiegreffe C, Simon R, Liu P, Juttner R, Crenshaw EB,3rd, Luyten FP, Jenkins NA, Copeland NG, Birchmeier C, Britsch S. Bcl11a is required for neuronal morphogenesis and sensory circuit formation in dorsal spinal cord development. *Development* 2012;**139**:1831-1841.

[39] Uetani N, Chagnon MJ, Kennedy TE, Iwakura Y, Tremblay ML. Mammalian motoneuron axon targeting requires receptor protein tyrosine phosphatases sigma and delta. *J Neurosci* 2006;**26**:5872-5880.

[40] Veeriah S, Brennan C, Meng S, Singh B, Fagin JA, Solit DB, Paty PB, Rohle D, Vivanco I, Chmielecki J, Pao W, Ladanyi M, Gerald WL, Liau L, Cloughesy TC, Mischel PS, Sander C, Taylor B, Schultz N, Major J, Heguy A, Fang F, Mellinghoff IK, Chan TA. The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated and mutated in glioblastoma and other human cancers. *Proc Natl Acad Sci U S A* 2009;**106**:9435-9440.

[41] Tsuritani K, Takeda J, Sakagami J, Ishii A, Eriksson T, Hara T, Ishibashi H, Koshihara Y, Yamada K, Yoneda Y. Cytokine receptor-like factor 1 is highly expressed in damaged human knee osteoarthritic cartilage and involved in osteoarthritis downstream of TGF-beta. *Calcif Tissue Int* 2010;**86**:47-57.

[42] Stenberg J, de Windt TS, Synnergren J, Hynsjo L, van der Lee J, Saris DB, Brittberg M, Peterson L, Lindahl A. Clinical Outcome 3 Years After Autologous Chondrocyte Implantation Does Not Correlate With the Expression of a Predefined Gene Marker Set in Chondrocytes Prior to Implantation but Is Associated With Critical Signaling Pathways. *Orthop J Sports Med* 2014;**2**:2325967114550781.

[43] Tartaglia M, Kalidas K, Shaw A, Song X, Musat DL, van der Burgt I, Brunner HG, Bertola DR, Crosby A, Ion A, Kucherlapati RS, Jeffery S, Patton MA, Gelb BD. PTPN11 mutations in Noonan syndrome: molecular spectrum, genotype-phenotype correlation, and phenotypic heterogeneity. *Am J Hum Genet* 2002;**70**:1555-1563.

[44] Määttä JH, Wadge S, MacGregor A, Karppinen J, Williams FM. ISSLS Prize Winner: Vertebral Endplate (Modic) Change is an Independent Risk Factor for Episodes of Severe and Disabling Low Back Pain. *Spine (Phila Pa 1976)* 2015;**40**:1187-1193.

[45] Kononenko O, Mityakina I, Galatenko V, Watanabe H, Bazov I, Gerashchenko A, Sarkisyan D, Iatsyshyna A, Yakovleva T, Tonevitsky A, Marklund N, Ossipov MH, Bakalkin G. Differential effects of left and right neuropathy on opioid gene expression in lumbar spinal cord. *Brain Res* 2018;**1695**:78-83.

[46] Wu L, Zhang S, Shkhyan R, Lee S, Gullo F, Eliasberg CD, Petrigliano FA, Ba K, Wang J, Lin Y, Evseenko D. Kappa opioid receptor signaling protects cartilage tissue against posttraumatic degeneration. *JCI Insight* 2017;**2**:88553.

[47] Bendix T, Sorensen JS, Henriksson GA, Bolstad JE, Narvestad EK, Jensen TS. Lumbar modic changes-a comparison between findings at low- and high-field magnetic resonance imaging. *Spine (Phila Pa 1976)* 2012;**37**:1756-1762.

FIGURE LEGENDS

Figure 1. Manhattan plot for the meta-analysis of GWAS of Modic changes in Northern European populations (Finland and UK, total N = 1829). Manhattan plot depicts p-values (-log10) presented on the y-axis and chromosomal position on the x-axis. Solid line depicts genome-wide significance (5e-8); dashed line depicts suggestive significance (5e-7).

Figure 2. Quantile-quantile plot for the meta-analysis of GWAS of Modic changes in Northern European populations (Finland and UK, total N = 1829). The Q-Q plot depicts the observed p-values against the expected p-values on -log10-scale. The test-statistic inflation (lambda) was 1.02, suggesting absence of population stratification in the dataset.

Figure 3. **Regional association plot of *PTPRD* locus.** The association p-value (-log10) is represented on the left-hand y-axis, recombination rate is displayed on the right-hand y-axis. Genomic location is shown on the x-axis (GRCh37/hg19). RefSeq genes are indicated in the bottom panel. Linkage disequilibrium r2 relative to the index single nucleotide variant rs1934268 (marked with diamond) is shown using the colors in the figure legend.

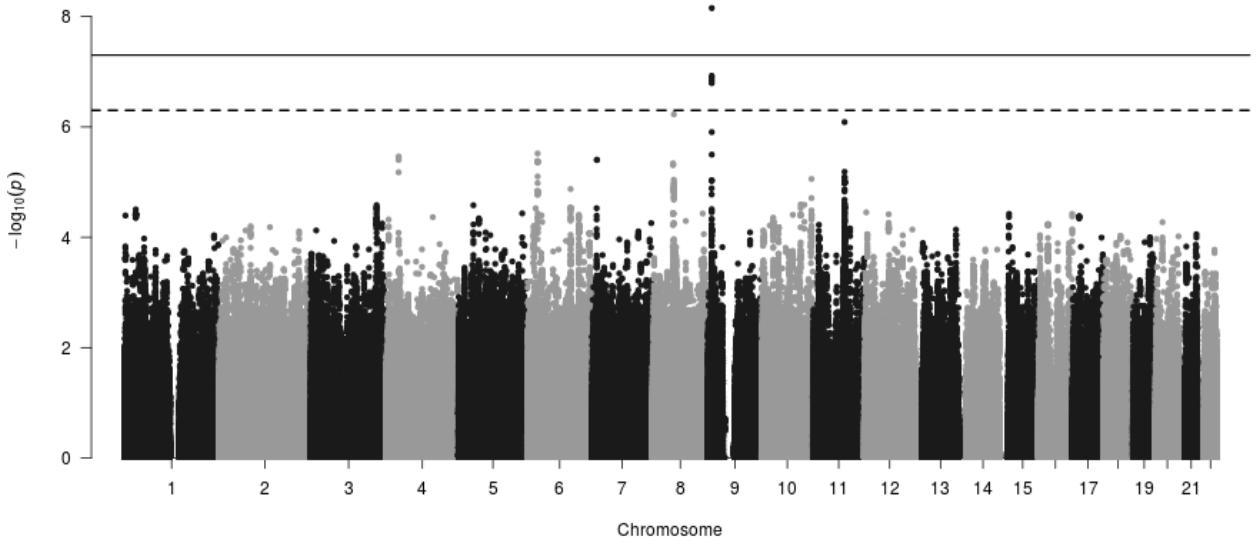

Figure 1

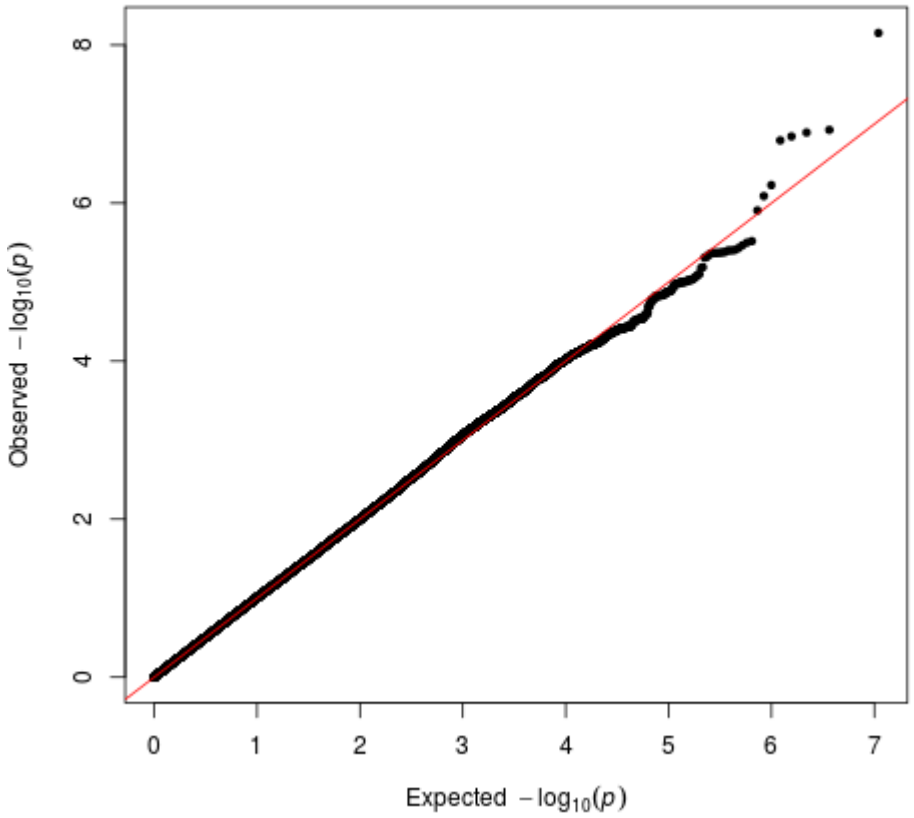

Figure 2

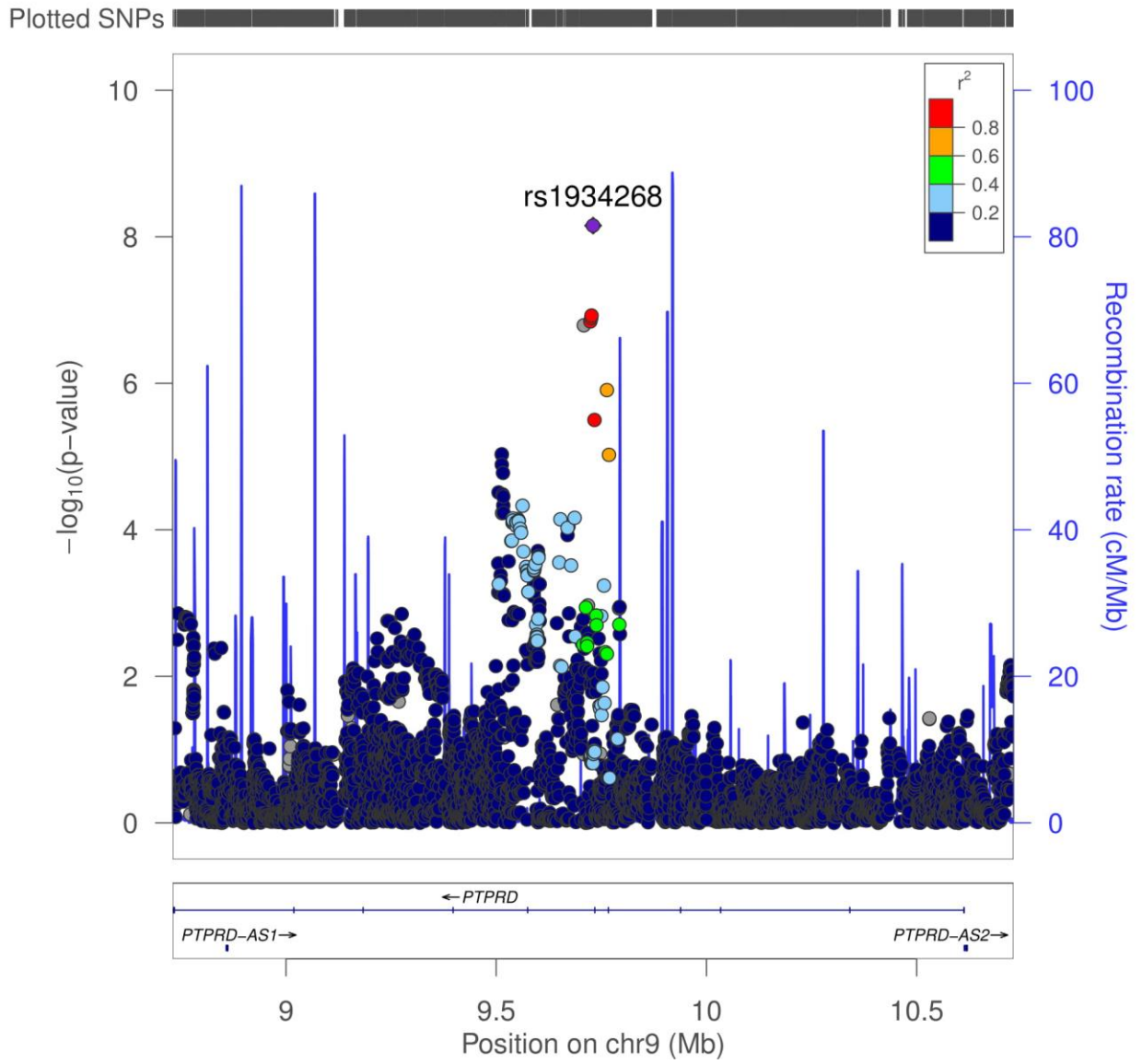
Plotted SNPs
rs1934268
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
PTPRD
PTPRD−AS1
PTPRD−AS2
Position on chr9 (Mb)

Figure 3

Supplementary table - The results of meta-GWAS for Modic changes in TwinsUK and NFBC1966 cohorts for SNPs with p-value 5e-7 < P < 1e-5

SNP ID	Chr:Pos	Effect allele	Other allele	EAF ± SE	Beta ± SE	OR (95% CI)	P-value	Direction	I^2	Het. P-value	Overlapped gene	Nearest gene (upstream or downstream)
rs36160095	4:32845987	A	G	0.938 ± 0.003	0.156 ± 0.034	1.87 (1.43-2.45)	3.4E-06	++	0	0.510	-	-
rs140476053	4:32843085	C	G	0.938 ± 0.003	0.155 ± 0.034	1.87 (1.43-2.44)	3.7E-06	++	0	0.522	-	-
rs143907932	4:32847206	T	C	0.062 ± 0.002	-0.155 ± 0.034	0.54 (0.41-0.7)	4.0E-06	--	0	0.528	-	-
rs150196231	4:32844355	T	C	0.063 ± 0.004	-0.151 ± 0.034	0.54 (0.42-0.71)	6.6E-06	--	0	0.634	-	-
rs201475493	6:27455134	A	AAATG	0.073 ± 0.010	-0.143 ± 0.031	0.56 (0.44-0.72)	3.0E-06	--	0	0.762	-	*ZNF184*
rs150644047	6:27457523	T	C	0.073 ± 0.010	-0.141 ± 0.031	0.57 (0.44-0.72)	4.2E-06	--	0	0.806	-	*ZNF184*
rs16867808	6:27463684	T	C	0.928 ± 0.010	0.142 ± 0.031	1.77 (1.39-2.26)	4.2E-06	++	0	0.807	-	*ZNF184*
rs78950897	6:27471298	T	G	0.072 ± 0.010	-0.142 ± 0.031	0.56 (0.44-0.72)	4.3E-06	--	0	0.809	-	*ZNF184*
rs74293985	6:27472488	T	C	0.928 ± 0.010	0.142 ± 0.031	1.77 (1.39-2.26)	4.3E-06	++	0	0.809	-	*ZNF184*
rs77052921	6:27473455	A	C	0.928 ± 0.010	0.142 ± 0.031	1.77 (1.39-2.26)	4.3E-06	++	0	0.810	-	*ZNF184*
rs77380993	6:27474646	T	C	0.928 ± 0.010	0.142 ± 0.031	1.77 (1.39-2.26)	4.3E-06	++	0	0.810	-	*ZNF184*
rs80183551	6:27454446	A	G	0.928 ± 0.010	0.141 ± 0.031	1.76 (1.38-2.25)	4.4E-06	++	0	0.820	-	*ZNF184*
rs76458558	6:27480547	A	G	0.928 ± 0.010	0.138 ± 0.031	1.74 (1.36-2.23)	7.9E-06	++	0	0.997	-	*ZNF184*
rs1017404	7:12695507	A	G	0.783 ± 0.015	0.088 ± 0.019	1.43 (1.23-1.66)	3.9E-06	++	53.2	0.144	-	-
rs1036372	7:12695169	A	G	0.783 ± 0.015	0.088 ± 0.019	1.43 (1.23-1.66)	4.0E-06	++	55.2	0.135	-	-
rs6473810	8:54244458	A	G	0.682 ± 0.002	-0.078 ± 0.017	0.73 (0.64-0.84)	4.6E-06	--	0	0.731	-	*OPRK1*
rs7831674	8:54224521	A	T	0.683 ± 0.001	-0.078 ± 0.017	0.73 (0.64-0.84)	4.9E-06	--	0	0.703	-	*OPRK1*
rs10107460	8:54247311	A	G	0.681 ± 0.003	-0.077 ± 0.017	0.73 (0.64-0.84)	4.9E-06	--	0	0.716	-	*OPRK1*
rs2658912	8:56280631	A	G	0.127 ± 0.015	-0.106 ± 0.024	0.65 (0.54-0.79)	9.1E-06	--	0	0.869	*XKR4*	-
rs1174582	9:9763539	C	G	0.821 ± 0.037	0.104 ± 0.022	1.52 (1.28-1.81)	1.2E-06	++	0	0.552	*PTPRD*	-
rs2761762	9:9733974	A	G	0.768 ± 0.044	0.089 ± 0.019	1.43 (1.23-1.66)	3.2E-06	++	0	0.337	*PTPRD*	-
rs4742607	9:9513109	A	G	0.462 ± 0.043	-0.071 ± 0.016	0.75 (0.66-0.85)	9.3E-06	--	0	0.635	*PTPRD*	-
rs4742606	9:9513039	A	G	0.462 ± 0.043	-0.071 ± 0.016	0.75 (0.66-0.85)	9.4E-06	--	0	0.634	*PTPRD*	-
rs1174587	9:9767971	A	G	0.829 ± 0.036	0.097 ± 0.022	1.48 (1.24-1.76)	9. 5E-06	++	0	0.524	*PTPRD*	-
rs74778840	10:131489527	C	G	0.902 ± 0.039	-0.131 ± 0.030	0.59 (0.47-0.75)	8.8E-06	--	0	0.792	*MGMT*	-

rs117285296	11:83055365	A	G	0.153 ± 0.028	0.110 ± 0.022	1.56 (1.31-1.85)	8.2E-07	++	0	0.329	-	-
rs2043208	11:83184839	A	G	0.235 ± 0.029	0.084 ± 0.019	1.4 (1.21-1.63)	6.5E-06	++	0	0.913	*DLG2*	-
rs9787813	11:83184256	A	C	0.229 ± 0.030	0.084 ± 0.019	1.4 (1.21-1.63)	8.1E-06	++	0	0.783	*DLG2*	-
rs183727937	11:83178853	T	C	0.650 ± 0.038	-0.076 ± 0.017	0.74 (0.64-0.84)	8.4E-06	--	0	0.409	*DLG2*	-
rs57023373	11:83181164	T	G	0.319 ± 0.028	0.076 ± 0.017	1.36 (1.19-1.55)	9.5E-06	++	0	0.492	*DLG2*	-
rs143569491	11:83195804	T	TG	0.230 ± 0.031	0.084 ± 0.019	1.4 (1.21-1.63)	9.8E-06	++	0	0.864	*DLG2*	-
rs17461216	11:83193813	A	G	0.230 ± 0.031	0.084 ± 0.019	1.4 (1.21-1.63)	9.8E-06	++	0	0.864	*DLG2*	-
rs1432050	11:83201563	T	C	0.229 ± 0.032	0.084 ± 0.019	1.4 (1.21-1.63)	9.9E-06	++	0	0.873	*DLG2*	-

Table legend

Genome-wide association study for Modic changes was carried out separately in TwinsUK and NFBC1966 cohorts followed by meta-analysis using inverse-variance weighting approach. Chr:pos, chromosome:position; EAF, effect allele frequency; Beta, coefficient of linear regression for effect allele vs other allele; OR, Odds ratio calculated using the following approximation: $OR = \exp(Beta/(\mu \times (1-\mu))$, where μ is the prevalence of MC (0.54 in the merged dataset); I^2, heterogeneity statistics.